LINEAR ALGEBRA AND ITS APPLICATIONS

ELSEVIER Linear Algebra and its Applications 288 (1999) 149–173

Sub-direct sums and positivity classes of matrices

Shaun M. Fallat *,[1], Charles R. Johnson [2]

Department of Mathematics, College of William and Mary, Williamsburg, VA 23187-8795, USA

Received 8 July 1997; accepted 14 September 1998

Submitted by W. Barrett

Abstract

It is well known that a direct sum is positive semidefinite if and only if each of the direct summands is positive semidefinite. In fact, it is also known that this statement remains true if positive semidefinite is replaced with: doubly nonnegative, completely positive, totally nonnegative, M-matrix and P-matrix, etc. For each of these classes we consider corresponding questions for a more general "sum" of two matrices, of which the direct sum and ordinary sum are special cases. © 1999 Elsevier Science Inc. All rights reserved.

1. Introduction

Let $0 \leqslant k \leqslant m, n$ and suppose that

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \in M_m(\mathbb{C}) \quad \text{and} \quad B = \begin{bmatrix} B_{22} & B_{23} \\ B_{32} & B_{33} \end{bmatrix} \in M_n(\mathbb{C}),$$

in which $A_{22}, B_{22} \in M_k(\mathbb{C})$. Then we call

* Corresponding author. E-mail: sfallat@math.wm.edu.

[1] The work of this author was supported by an NSERC PGS B award.

[2] The work of this author was supported in part by the Office of Naval Research contract N00014-90-J-1739 and NSF grant 92-00899.

PII: S0024-3795(98)10194-5

$$C = \begin{bmatrix} A_{11} & A_{12} & 0 \\ A_{21} & A_{22}+B_{22} & B_{23} \\ 0 & B_{32} & B_{33} \end{bmatrix}$$

the *k-subdirect sum* of A and B, which we denote by $A \oplus_k B$. When the value of k is irrelevant or clear, we may just refer to a *subdirect sum*, and, when $k = 0$, a 0-subdirect sum is a familiar direct sum, and $\oplus_0$ will be abbreviated to $\oplus$. Of course, the k-subdirect sum of two matrices is generally not commutative. We are interested in the following key positivity classes of matrices: positive (semi)-definite (PD, (PSD)); P-matrices (positive principal minors) (P); P_0-matrices (nonnegative principal minors) (P_0); M-matrices (nonpositive off-diagonal entries and positive principal minors) (M); totally nonnegative matrices (all minors nonnegative) (TN); completely positive matrices (matrices of the form BB^{T} with B entry-wise nonnegative) (CP); doubly nonnegative matrices (positive semidefinite and entry-wise nonnegative) (DN); and symmetric M-matrices (SM). In each case, it is elementary that a direct sum lies in the class if and only if each direct summand lies in the class. We are interested in corresponding questions for subdirect sums, and it is appropriate to consider 1-subdirect sums and k-subdirect sums ($k > 1$) separately. The notion of subdirect sums arise naturally in a variety of ways, such as matrix completion problems [1–4] and in understanding the structure of classes, for example [5]. In particular, the special structure of a 1-subdirect sum of two matrices has appeared in Refs. [1,4]. We expect that they will continue to arise in new ways.

For each of the classes: PD, PSD, M, TN, P, P_0, CP, DN and SM, we address four natural questions: (I) If A and B lie in the class must a 1-subdirect sum C lie in the class; (II) If

$$C = \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{21} & C_{22} & C_{23} \\ 0 & C_{32} & C_{33} \end{bmatrix}$$

lies in the class, may C be written as $C = A \oplus_1 B$, such that A and B lie in the class when C_{22} is 1×1; and (III) and (IV) the corresponding questions with 1 replaced by $k > 1$.

2. Preliminary results

For $A \in M_n(\mathbb{C})$, we let det A denote the determinant of A, and let I_n denote the $n \times n$ identity matrix. For $A \in M_n(\mathbb{C})$, $\alpha, \beta \subseteq N \equiv \{1, 2, \ldots, n\}$, the submatrix of A lying in the rows indexed by α and the columns indexed by β will be denoted $A[\alpha|\beta]$. Similarly, the submatrix obtained from A by deleting the rows indexed by α and the columns indexed by β will be denoted $A(\alpha|\beta)$. If $\alpha = \beta$, then the principal submatrix $A[\alpha|\alpha]$ ($A(\alpha|\alpha)$) is abbreviated to $A[\alpha]$ ($A(\alpha)$). In the special case in which $\alpha = \{i\}$, we denote $A[\alpha]$ and $A(\alpha)$ by $A[i]$ and $A(i)$, respectively. If x is a row (column) n-vector, then $x[\alpha]$ denotes, the subvector of x retaining columns (rows) indexed by α. As usual, $|\alpha|$ denotes the cardinality of α.

We begin with a simple determinantal identity, that is useful in consideration of 1-subdirect sums. We provide a proof here for completeness.

Lemma 2.1. *Let*

$$A = \begin{bmatrix} A_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in M_m(\mathbb{C}) \quad \textit{and} \quad B = \begin{bmatrix} b_{22} & b_{23} \\ b_{32} & B_{33} \end{bmatrix} \in M_n(\mathbb{C}),$$

in which a_{22} and b_{22} are both 1×1. Then

$$\det(A \oplus_1 B) = \det A_{11} \det B + \det A \det B_{33}. \tag{1}$$

Proof. It is routine to verify that if A_{11} and B_{33} are both singular, then $A \oplus_1 B$ is necessarily singular; hence Eq. (1) holds in this special case. Suppose, then, without loss of generality, that A_{11} is nonsingular. Using Schur complements (see, for example, [6], p. 22) it follows that

$$\det \begin{bmatrix} A_{11} & a_{12} & 0 \\ a_{21} & a_{22} + b_{22} & b_{23} \\ 0 & b_{32} & B_{33} \end{bmatrix} = \det A_{11} \det \begin{bmatrix} a_{22} + b_{22} - a_{21}A_{11}^{-1}a_{12} & b_{23} \\ b_{32} & B_{33} \end{bmatrix}.$$

Expanding the latter determinant by the first row gives

$$\begin{aligned} &\det A_{11} \det \begin{bmatrix} b_{22} & b_{23} \\ b_{32} & B_{33} \end{bmatrix} + \det A_{11}(a_{22} - a_{21}A_{11}^{-1}a_{12}) \det B_{33} \\ &= \det A_{11} \det B + \det A \det B_{33}. \end{aligned}$$

The affirmative answer to questions (I)–(IV) for positive semidefinite matrices has been known for some time, e.g. [5]. This observation is summarized in the following theorem.

Theorem 2.2 ([5], Lemma 2). *Let*

$$C = \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{12}^* & C_{22} & C_{23} \\ 0 & C_{23}^* & C_{33} \end{bmatrix}$$

be an $n \times n$ Hermitian matrix with C_{11}, C_{22} and C_{33} square. Then C is positive semidefinite if and only if there exist Hermitian matrices A_{22} and B_{22} such that $C_{22} = A_{22} + B_{22}$ and such that

$$A = \begin{bmatrix} C_{11} & C_{12} \\ C_{12}^* & A_{22} \end{bmatrix} \quad \textit{and} \quad B = \begin{bmatrix} B_{22} & C_{23} \\ C_{23}^* & C_{33} \end{bmatrix}$$

are positive semidefinite.

Using the notation developed here, Theorem 2.2 is equivalent to the following, C (as in Theorem 2.2) is positive semidefinite if and only if C may be written as $C = A \oplus_k B$ (k is the order of C_{22}), with A and B positive semidefinite. Thus the above theorem gives affirmative answers to questions (I)–(IV) for the class of positive semidefinite matrices.

Also considered in Ref. [5] were questions (I) and (II) for the classes DN and CP. We present their results for these classes in the following proposition.

Proposition 2.3 ([5], Cors. 1,2). *Let*

$$C = \begin{bmatrix} C_{11} & c_{12} & 0 \\ c_{12}^* & c_{22} & c_{23} \\ 0 & c_{23}^* & C_{33} \end{bmatrix}$$

be an $n \times n$ Hermitian matrix with C_{11} and C_{33} square, and with c_{22} 1×1. Then C is DN (CP) if and only if c_{22} may be written as $c_{22} = a_{22} + b_{22}$ so that

$$A = \begin{bmatrix} C_{11} & c_{12} \\ c_{12}^* & a_{22} \end{bmatrix} \quad \textit{and} \quad B = \begin{bmatrix} b_{22} & c_{23} \\ c_{23}^* & C_{33} \end{bmatrix}$$

are both DN (CP).

It is well known, see, for example, [5] that, if $n \leqslant 4$, then CP = DN, and for $n \geqslant 5$, CP is properly contained in DN. This fact will be used in Section 3. It is useful to note that the each of the classes PSD, PD, DN and CP is closed under addition. Recall that a Z-matrix is a real matrix with nonpositive off-diagonal entries. Thus A is an M-matrix if and only if A is a Z- and a P-matrix. A matrix is SM if and only if it is a positive definite Z-matrix. Since PD and Z matrices are each closed under addition, the class SM is also closed under addition. It is easy to verify that the remaining classes M, TN, P and P_0, are not closed under addition, when $n > 1$. For any matrix A in any of the positivity classes discussed in the introduction, it follows that each principal submatrix of A remains in the same positivity class. In the special class TN, all submatrices (principal and nonprincipal) are also TN. These inheritance properties will be useful later.

As M-matrices are considered within, we outline selected facts about M-matrices and symmetric M-matrices; for more details, see, for example, ([7], p. 112). If A is an M-matrix, then there exists an entry-wise positive diagonal matrix D such that AD is stricly row diagonally dominant. If A is a symmetric

M-matrix, then there exists an entry-wise positive diagonal matrix D such that DAD^{T} is stricly row and column diagonally dominant.

The proof we present for determining the answer to question (IV), for the class P, requires a few preliminary results which we state here. For other results concerning P-matrices, see Ref. [7]. The next result is elementary and we state it without proof.

Lemma 2.4. *Let $A \in M_2(\mathbb{R})$, and let S be any fixed nonzero 2×2 real skew-symmetric matrix. Then there exists a $T > 0$ so that, for all $t \geqslant T$,* $\det(A + tS) > 0$.

We now extend this perturbation result to a larger class of matrices.

Lemma 2.5. *Let $A \in M_n(\mathbb{R})$, and suppose that A is permutationally similar to a matrix (which we denote by A) of the form,*

$$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix},$$

in which A_4 is 2×2, and $A(n)$ and $A(n-1)$ are P-matrices. Then for any fixed nonzero 2×2 real skew-symmetric matrix S, there exists a $T > 0$ so that, for all $t \geqslant T$,

$$B_t = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 + tS \end{bmatrix},$$

is a P-matrix.

Proof. Let S be as indicated, and let B_t be defined as above. Let $\hat{N} = \{\alpha \subseteq N : n \in \alpha \text{ and } n-1 \in \alpha\}$. To show that B_t is a P-matrix it is enough to show $\det(B_t[\alpha]) > 0$, for all $\alpha \in \hat{N}$, the other principal minors being positive by assumption. Let $\alpha \in \hat{N}$ and let

$$B_t[\alpha] = \begin{bmatrix} B_1 & B_2 \\ B_3 & A_4 + tS \end{bmatrix},$$

in which $t > 0$. Then

$$\det(B_t[\alpha]) = \det B_1 \cdot \det([B_t[\alpha]/B_1] + tS),$$

where $B_t[\alpha]/B_1 = A_4 - B_3 B_1^{-1} B_2$. By Lemma 2.4 choose $T_\alpha > 0$ so that $\det(B_t[\alpha]/B_1 + tS) > 0$, for all $t \geqslant T_\alpha$. Let $T = \max_{\alpha \in \hat{N}}\{T_\alpha\}$. Then, for all $t \geqslant T$, B_t is a P-matrix. $\square$

Lemma 2.6. *Let $A \in M_n(\mathbb{R})$, and suppose that A is permutationally similar to a matrix (which we denote by A) of the form,*

$$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix},$$

in which A_4 is $k \times k$ ($k \geqslant 2$). Also suppose $A(n)$ and $A(\{n-k+1,\ldots,n-1\})$ are P-matrices. Then for any fixed $k \times k$ real skew-symmetric matrix S with all off-diagonal entries nonzero in the last row and column and all other off-diagonal entries zero, there exists a $T > 0$ so that, for all $t \geqslant T$,

$$B_t = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 + tS \end{bmatrix},$$

is a P-matrix.

Proof. The proof will use induction on r, the size of A_4. By Lemma 2.5 the result holds if $r = 2$. Assume the result holds for $2 \leqslant r \leqslant k-1$. Let A, S and B_t be defined as above. Fix $\alpha = \{1,2,\ldots,n-k,n-k+1,n\}$. Then

$$C_t \equiv B_t[\alpha] = \begin{bmatrix} A_1 & B_2 \\ B_3 & B_4 + tS' \end{bmatrix},$$

where B_4 and S' are 2×2 and $S' = S[\{1,k\}]$. Furthermore, by assumption $C_t(n-k+1)$ and $C_t(n-k+2)$ are P-matrices. Then by Lemma 2.5 there exists a $T' > 0$ such that, for all $t \geqslant T'$, C_t is a P-matrix. Define $\bar{S}$ to be the nonzero $k \times k$ skew-symmetric matrix with only the $(1,k)$ and $(k,1)$ entries nonzero and equal to s_{1k} and $-s_{1k}$, respectively. Then

$$\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 + T'\bar{S} \end{bmatrix}$$

satisfies the hypotheses of the lemma with $r = k-1$. It follows by induction that there exists $T > 0$ (here T is chosen large enough) such that for all $t \geqslant T$, B_t is a P-matrix. □

The above sequence of lemmata may be used to give another proof of the fact that any *partial combinatorially symmetric P-matrix*, can be completed to a combinatorially symmetric P-matrix. This result was first proved in Ref. [8].

In Section 3, we consider the classes (of the ones mentioned in the introduction), which are closed under addition, namely, PSD, PD, CP, DN and SM. For each of these classes we answer questions (I)–(IV). We divide the remaining four clasess into two sections. In Section 4, we consider questions (I)–(IV) for the classes M, P and P_0. In the final section only the class TN is considered. Finally, we provide a summary of our results via Table 1.

3. Classes closed under addition

In this section we answer questions (I)–(IV) for the classes that are closed under addition: PSD, PD, CP, DN and SM. When $k = n = m$, it is simple to give an affirmative answer to question (IV), for any of the above five classes; that is, any $n \times n$ matrix in one of PSD, PD, CP, DN or SM, can be written as the sum (n-subdirect sum) of two matrices in the same class. As already noted a direct sum (0-subdirect sum) is in any one of the above classes if and only if each summand is in the class. We consider questions (I)–(IV) for general k.

As stated in Section 2, Theorem 2.2 gives affirmative answers to questions (I)–(IV) for the class PSD. Moreover, Theorem 2.2 may be used to prove affirmative answers to questions (I)–(IV) for the class PD. We provide a proof for completeness.

Proposition 3.1. *Let*

$$C = \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{12}^* & C_{22} & C_{23} \\ 0 & C_{23}^* & C_{33} \end{bmatrix}$$

be an $n \times n$ Hermitian matrix with C_{11} and C_{33} square, and with C_{22} $k \times k$. Then C is positive definite if and only if C may be written as $C = A \oplus_k B$, in which both A and B are positive definite.

Proof. Suppose C is PD. Then there exists an $\alpha > 0$ so that $C' = C - \alpha I$ is PSD. By Theorem 2.2, C' may be written as $C' = A' \oplus_k B'$, in which A' and B' are both PSD. Let

$$A = A' + \begin{bmatrix} \alpha I & 0 \\ 0 & (\alpha/2)I_k \end{bmatrix}$$

and

$$B = B' + \begin{bmatrix} (\alpha/2)I_k & 0 \\ 0 & \alpha I \end{bmatrix}.$$

Then A and B are both PD, and

$$A \oplus_k B = C' + \alpha I = C.$$

Conversely, suppose A and B are both PD. Then there exist $\alpha > 0$ and $\beta > 0$ so that $A' = A - \alpha I$ and $B' = B - \beta I$ are both PSD. By Theorem 2.2, $A' \oplus_k B'$ is PSD. However, it is easy to verify that

$$A \oplus_k B = A' \oplus_k B' + D,$$

where

$$D = \begin{bmatrix} \alpha I & 0 & 0 \\ 0 & (\alpha + \beta)I_k & 0 \\ 0 & 0 & \beta I \end{bmatrix}.$$

Thus it follows that $A \oplus_k B$ is PD. $\square$

We note that Theorem 2.2 may also be proven from this proposition, which may itself be proven independently.

Suppose $A \in M_n(\mathbb{C})$ and $B \in M_m(\mathbb{C})$. Then for a given k with $1 \leqslant k \leqslant \min\{m, n\}$, let

$$A' = A \oplus 0_{m-k} \quad \text{and} \quad B' = 0_{n-k} \oplus B,$$

in which 0_l is the $l \times l$ zero matrix. It is then clear that,

$$A \oplus_k B = A' + B'.$$

We are now in a position to prove the next proposition.

Proposition 3.2. *Suppose $\mathscr{S}$ is any one of CP, DN or SM. Let $A \in M_n(\mathbb{R})$ and $B \in M_m(\mathbb{R})$. If A and B are in $\mathscr{S}$, then $A \oplus_k B \in \mathscr{S}$, for each k, $1 \leqslant k \leqslant \min\{m, n\}$.*

Proof. Consider the class CP. Since A and B are CP, $A' = A \oplus 0_{m-k}$ and $B' = B \oplus 0_{n-k}$ are both CP. Hence $A' + B'$ is CP, since CP is closed under addition. However, by the observation preceeding this proposition, $A \oplus_k B = A' + B'$. Thus $A \oplus_k B$ is CP. The argument for the class DN is similar. Finally, consider the class SM, and suppose A and B are both SM. Then, since A and B are PD $A \oplus_k B$ is PD, by Proposition 3.1. Furthermore, $A \oplus_k B$ has nonpositive off-diagonal entries. Hence $A \oplus_k B$ is SM. This completes the proof. $\square$

For the two special classes, CP and SM the converse to Proposition 3.2 holds, and hence yields an affirmitive answer to question (III).

Proposition 3.3. *Let*

$$C = \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{12}^{\mathrm{T}} & C_{22} & C_{23} \\ 0 & C_{23}^{\mathrm{T}} & C_{33} \end{bmatrix} \in M_n(\mathbb{R})$$

be an $n \times n$ symmetric matrix with C_{11} and C_{33} square, and with C_{22} $k \times k$. Suppose C is CP. Then C may be written as $C = A \oplus_k B$, in which both A and B are CP.

Proof. Since C is CP, $C = RR^{\mathrm{T}}$, in which R is an $n \times m$ entry-wise nonnegative matrix. Equivalently, $C = \sum_{i=1}^{m} r_i r_i^{\mathrm{T}}$, where r_i is the ith column of R. For simplicity of notation, let $\alpha, \beta \subseteq N$ be such that

$$C[\alpha] = \begin{bmatrix} C_{11} & C_{12} \\ C_{12}^{T} & C_{22} \end{bmatrix} \quad \text{and} \quad C[\beta] = \begin{bmatrix} C_{22} & C_{23} \\ C_{23}^{T} & C_{33} \end{bmatrix}.$$

Since each r_i is nonnegative and because of the form of C, all of the nonzero entries of an r_i either lie wholly in α or wholly in β. Let A' be the sum of the $r_i r_i^{\mathrm{T}}$ over those i for which the nonzero entries of r_i lie only in α. Similarly, let B' be the sum of the $r_i r_i^{\mathrm{T}}$ over those i for which the nonzero entries of r_i lie only in β. Thus A' and B' are CP. Let $A = A'[\alpha]$ and $B = B'[\beta]$. Then A and B are CP and since $A' + B' = C$, it follows that $A \oplus_k B = C$. □

Proposition 3.4. *Let*

$$C = \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{12}^{\mathrm{T}} & C_{22} & C_{23} \\ 0 & C_{23}^{\mathrm{T}} & C_{33} \end{bmatrix} \in M_n(\mathbb{R})$$

be an $n \times n$ symmetric matrix with C_{11} and C_{33} square, and with C_{22} $k \times k$. Suppose C is SM. Then C may be written as $C = A \oplus_k B$, in which both A and B are SM.

Proof. Suppose C is a symmetric M-matrix. Let e denote the vector (whose number of components is determined by context), consisting of all ones. Let D be an entry-wise positive diagonal matrix such that CD is strictly row diagonally dominant (see, for example, [7], p. 114). Since CD has positive main diagonal and nonpositive off-diagonal entries it follows that CD is strictly row diagonally dominant if and only if $CDe > 0$. Since C is symmetric it follows by direct calculation that DCD satisfies $DCDe > 0$ and $e^{\mathrm{T}}DCD > 0$, that is, DCD is both stricly row and column dominant. Partition D conformally with C. Then

$$DCD = \begin{bmatrix} D_{11}C_{11}D_{11} & D_{11}C_{12}D_{22} & 0 \\ (D_{11}C_{12}D_{22})^{\mathrm{T}} & D_{22}C_{22}D_{22} & D_{22}C_{23}D_{33} \\ 0 & (D_{22}C_{23}D_{33})^{\mathrm{T}} & D_{33}C_{33}D_{33} \end{bmatrix}.$$

Let $D_{22}C_{22}D_{22} = S_{22} + T_{22}$, in which S_{22} is diagonal and T_{22} has zero main diagonal. Since DCD is row diagonally dominant and DCD is SM,

$$S_{22}e + (D_{11}C_{12}D_{22})^{\mathrm{T}}e + T_{22}e + (D_{22}C_{23}D_{33})e > 0,$$

where the inequality above is entry-wise. Therefore,

$$S_{22}e + ((D_{11}C_{12}D_{22})^{\mathrm{T}} + \tfrac{1}{2}T_{22})e + (\tfrac{1}{2}T_{22} + D_{22}C_{23}D_{33})e > 0.$$

Hence S_{22} may be written as $S_{22} = S' + S''$, such that

$$S'e + ((D_{11}C_{12}D_{22})^{\mathrm{T}} + \tfrac{1}{2}T_{22})e > 0,$$

and

$$S''e + (\tfrac{1}{2}T_{22} + D_{22}C_{23}D_{33})e > 0.$$

Let $A'_{22} = S' + \frac{1}{2}T_{22}$ and $B'_{22} = S'' + \frac{1}{2}T_{22}$. Then

$$A' = \begin{bmatrix} D_{11}C_{11}D_{11} & D_{11}C_{12}D_{22} \\ (D_{11}C_{12}D_{22})^{\mathrm{T}} & A'_{22} \end{bmatrix}$$

and

$$B' = \begin{bmatrix} B'_{22} & D_{22}C_{23}D_{33} \\ (D_{22}C_{23}D_{33})^{\mathrm{T}} & D_{33}C_{33}D_{33} \end{bmatrix},$$

are both symmetric M-matrices. Furthermore,

$$A' = \begin{bmatrix} D_{11} & 0 \\ 0 & D_{22} \end{bmatrix} \cdot \begin{bmatrix} C_{11} & C_{12} \\ C_{12}^{\mathrm{T}} & D_{22}^{-1}A'_{22}D_{22}^{-1} \end{bmatrix} \cdot \begin{bmatrix} D_{11} & 0 \\ 0 & D_{22} \end{bmatrix},$$

and

$$B' = \begin{bmatrix} D_{22} & 0 \\ 0 & D_{33} \end{bmatrix} \cdot \begin{bmatrix} D_{22}^{-1}B'_{22}D_{22}^{-1} & C_{23} \\ C_{23}^{\mathrm{T}} & C_{33} \end{bmatrix} \cdot \begin{bmatrix} D_{22} & 0 \\ 0 & D_{33} \end{bmatrix}.$$

Since $C_{22} = D_{22}^{-1}(A'_{22} + B'_{22})D_{22}^{-1}$, it follows that C may be written as $C = A \oplus_k B$, where A and B are both symmetric M-matrices. □

Finally, consider the class DN. For the special case when $k = 1$, Proposition 2.3 gives an affirmative answer to question (II). However, when $k \geqslant 2$, the answer to question (IV) is negative, in general.

Example 3.5. Let

$$C = \begin{bmatrix} 4 & 2 & 2 & 0 & 0 \\ 2 & 4 & 0 & 0 & 2 \\ 2 & 0 & 4 & 3 & 0 \\ 0 & 0 & 3 & 4 & 2 \\ 0 & 2 & 0 & 2 & 4 \end{bmatrix}.$$

In Ref. [5] it is noted that C is DN but not CP. Recall that $n = 5$ is the smallest n for which there exists a matrix that is DN but not CP. Here C can be considered either as a 2 or 3-subdirect sum. Suppose that C could be written as $C = A \oplus_3 B$ (the argument for a 2-subdirect sum is similar), with both A and B in DN and 4×4. Since A and B are 4×4, they are CP. Let $\hat{A}$ ($\hat{B}$) be the matrix obtained from A (B) by bordering the last (first) row and column of A (B) by zeros. Then $\hat{A}$ and $\hat{B}$ are CP and since $C = \hat{A} + \hat{B}$, C must be CP. This is a contradiction, hence no such decomposition of C exists.

4. M, P and P_0-matrices

In this section we consider only three classes M, P and P_0. Note that these classes are not closed under addition. This will be used later. Our first result of the section gives an affirmative answer to question (IV) for the class M.

Proposition 4.1. *Let*

$$C = \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{21} & C_{22} & C_{23} \\ 0 & C_{32} & C_{33} \end{bmatrix} \in M_n(\mathbb{R})$$

be an $n \times n$ matrix with C_{11} and C_{33} square, and with C_{22} $k \times k$. Suppose C is M. Then C may be written as $C = A \oplus_k B$, in which both A and B are M.

Proof. Suppose C is an M-matrix. Let D be an entry-wise positive diagonal matrix such that CD is strictly row diagonally dominant (see Ref. [7], p. 114). Assume D is partitioned conformally with C. Then

$$CD = [r_{ij}] = \begin{bmatrix} C_{11}D_{11} & C_{12}D_{22} & 0 \\ C_{21}D_{11} & C_{22}D_{22} & C_{23}D_{33} \\ 0 & C_{32}D_{22} & C_{33}D_{33} \end{bmatrix}.$$

Let $C_{22}D_{22} = S_{22} + T_{22}$, in which S_{22} is diagonal and T_{22} has zero main diagonal. Further, let $T_{22} = L_{22} + U_{22}$, in which L_{22} (U_{22}) is strictly lower (upper) triangular. Let row i be any row of CD that involves the overlapping block $C_{22}D_{22}$. Since $r_{ii} > -\sum_{j\neq i} r_{ij}$, there exists $\varepsilon > 0$ such that $r_{ii} = -\sum_{j\neq i} r_{ij} + \varepsilon$. Choose $a_{ii} \in [0, r_{ii}]$ so that $a_{ii} = -\sum_{j<i} r_{ij} + \varepsilon/2$. Then $b_{ii} = r_{ii} - a_{ii}$ necessarily satisfies, $b_{ii} = -\sum_{j>i} r_{ij} + \varepsilon/2$. Let $A'_{22} = \mathrm{diag}(a_{ii}) + L_{22}$ and $B'_{22} = \mathrm{diag}(b_{ii}) + U_{22}$. Then

$$A' = \begin{bmatrix} C_{11}D_{11} & C_{12}D_{22} \\ C_{21}D_{11} & A'_{22} \end{bmatrix} \quad \text{and} \quad B' = \begin{bmatrix} B'_{22} & C_{23}D_{33} \\ C_{32}D_{22} & C_{33}D_{33} \end{bmatrix},$$

are both M-matrices. Furthermore,

$$A' = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & A'_{22}D_{22}^{-1} \end{bmatrix} \cdot \begin{bmatrix} D_{11} & 0 \\ 0 & D_{22} \end{bmatrix},$$

and

$$B' = \begin{bmatrix} B'_{22}D_{22}^{-1} & C_{23} \\ C_{32} & C_{33} \end{bmatrix} \cdot \begin{bmatrix} D_{22} & 0 \\ 0 & D_{33} \end{bmatrix}.$$

Since $C_{22} = A'_{22}D_{22}^{-1} + B'_{22}D_{22}^{-1}$, it follows that C can be written as $C = A \oplus_k B$, where A and B are both M-matrices (see Ref. [7], p. 127). $\square$

The converse to the above proposition does not hold, in general, for M, P nor P_0, when $k \geqslant 2$. Since these classes are not closed under addition, it would be trivial to construct examples of matrices A and B each in one of the classes, and have $A \oplus_k B$, $k \geqslant 2$, not in the class, simply by arranging that the sum in the overlapping block is not in the class and noting the inheritance property for principal submatrices. However, our example shows that even more can go wrong.

Example 4.2. Let

$$A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 5 & 0 \\ -1 & -9 & 5 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 5 & -9 & -1 \\ 0 & 5 & -1 \\ -1 & -1 & 2 \end{bmatrix}.$$

Then A and B are both M-matrices (and P, P_0-matrices). However,

$$A \oplus_2 B = \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 10 & -9 & -1 \\ -1 & -9 & 10 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix}$$

is not an M-matrix (or a P, P_0-matrix) since $\det(A \oplus_2 B) = -76$. Note that the sum in the overlapping positions is an M-matrix (P, P_0-matrix).

Consider now question (IV) for the class P. First we consider the special case when $k = 1$ (or questions (I) and (II)) separately. This result is used in the proof for $k > 1$.

Lemma 4.3. *Let*

$$C = \begin{bmatrix} C_{11} & c_{12} & 0 \\ c_{21} & c_{22} & c_{23} \\ 0 & c_{32} & C_{33} \end{bmatrix} \in M_n(\mathbb{R})$$

be an $n \times n$ matrix with C_{11} and C_{33} square, and with c_{22} 1×1. Then C is P if and only if C may be written as $C = A \oplus_1 B$, in which both A and B are P.

Proof. Suppose A and B are P-matrices. Then by Eq. (1) it follows that $\det(A \oplus_1 B) > 0$. It is straightforward to verify that the remaining principal minors of $A \oplus_1 B$ are either, a 1-subdirect sum of principal submatrices of A and B or a direct sum of principal submatrices of A and B. In either case, the minor is easily shown to be positive, hence $A \oplus_1 B$ is a P-matrix. Conversely, suppose C is a P-matrix and let $a_{22} \in (0, c_{22})$, and $b_{22} = c_{22} - a_{22}$. By Eq. (1) it follows that

$$\det C_{11} \det \begin{bmatrix} b_{22} & c_{23} \\ c_{32} & C_{33} \end{bmatrix} + \det C_{33} \det \begin{bmatrix} C_{11} & c_{12} \\ c_{21} & a_{22} \end{bmatrix} > 0.$$

Hence, using Schur complements,

$$\det C_{11} \det C_{33}(a_{22} + b_{22} - c_{21}C_{11}^{-1}c_{12} - c_{23}C_{33}^{-1}c_{32}) > 0,$$

which implies that $c_{22} - c_{23}C_{33}^{-1}c_{32} > c_{21}C_{11}^{-1}c_{12}$ (recall that $a_{22} + b_{22} = c_{22}$). Since C is a P-matrix we have

$$\max\{0, c_{21}C_{11}^{-1}c_{12}\} < \min\{c_{22}, c_{22} - c_{23}C_{33}^{-1}c_{32}\}.$$

Hence we may choose a_{22} so that,

$$\max\{0, c_{21}C_{11}^{-1}c_{12}\} < a_{22} < \min\{c_{22}, c_{22} - c_{23}C_{33}^{-1}c_{32}\}. \tag{2}$$

Let

$$A = \begin{bmatrix} C_{11} & c_{12} \\ c_{21} & a_{22} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{22} & c_{23} \\ c_{32} & C_{33} \end{bmatrix}.$$

If a_{22} satisfies the inequalities in Eq. (2), then $\det A > 0$ and $\det B > 0$. The idea to complete the proof is to apply this argument several times to obtain an interval of choices for a_{22} so that A and B are P-matrices. Suppose A is $m \times m$ and B is $r \times r$ $(m + r - 1 = n)$. Let $\hat{M} = \{\alpha \subseteq \{1, 2, \ldots, m\} : m \in \alpha$, and $|\alpha| \geqslant 2\}$, and let $\hat{R} = \{\beta \subseteq \{1, 2, \ldots, r\} : 1 \in \beta$, and $|\beta| \geqslant 2\}$. To show that A (B) is a P-matrix it is enough to show $\det A[\alpha] > 0$ ($\det B[\beta] > 0$), for all $\alpha \in \hat{M}$ ($\beta \in \hat{R}$). The other principal minors are positive by assumption. Let $\alpha \in \hat{M}$ be fixed (yet arbitrary), and let

$$A[\alpha] = \begin{bmatrix} A_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Choose any $\beta \in \hat{R}$, and let

$$B[\beta] = \begin{bmatrix} b_{22} & b_{23} \\ b_{32} & B_{33} \end{bmatrix}.$$

Then $A[\alpha] \oplus_1 B[\beta]$ is a principal submatrix of C and hence is a P-matrix. Thus applying a similar argument as above, we have,

$$\max\{0, a_{21}A_{11}^{-1}a_{12}\} < \min\{c_{22}, c_{22} - b_{23}B_{33}^{-1}b_{32}\}.$$

Since this inequality is satisfied for any $\alpha \in \hat{M}$ and $\beta \in \hat{R}$, it follows that

$$\max\{0, \max_{\alpha \in \hat{M}}\{a_{21}A_{11}^{-1}a_{12}\}\} < \min\{c_{22}, \min_{\beta \in \hat{R}}\{c_{22} - b_{23}B_{33}^{-1}b_{32}\}\}.$$

Hence we can choose a_{22} so that,

$$\max\{0, \max_{\alpha \in \hat{M}}\{a_{21}A_{11}^{-1}a_{12}\}\} < a_{22} < \min\{c_{22}, \min_{\beta \in \hat{R}}\{c_{22} - b_{23}B_{33}^{-1}b_{32}\}\}.$$

Then by construction, if

$$A = \begin{bmatrix} C_{11} & c_{12} \\ c_{21} & a_{22} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{22} & c_{23} \\ c_{32} & C_{33} \end{bmatrix},$$

then A and B are both P-matrices, and $A \oplus_1 B = C$. This completes the proof. □

Suppose A and B are M-matrices. Then A and B are P-matrices. Hence $A \oplus_1 B$ is a P-matrix, by Lemma 4.3. Since $A \oplus_1 B$ is a Z-matrix (off-diagonal entries nonpositive), it follows that $A \oplus_1 B$ is an M-matrix. This observation along with Proposition 4.1 prove the following.

Corollary 4.4. *Let*

$$C = \begin{bmatrix} C_{11} & c_{12} & 0 \\ c_{21} & c_{22} & c_{23} \\ 0 & c_{32} & C_{33} \end{bmatrix} \in M_n(\mathbb{R})$$

be an $n \times n$ matrix with C_{11} and C_{33} square, and with c_{22} 1×1. Then C is M if and only if C may be written as $C = A \oplus_1 B$, in which both A and B are M.

The next result gives an affirmative answer to question (IV) for the class P.

Theorem 4.5. *Let*

$$C = \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{21} & C_{22} & C_{23} \\ 0 & C_{32} & C_{33} \end{bmatrix} \in M_n(\mathbb{R}),$$

where C_{11} and C_{33} are square, and with C_{22} $k \times k$. If C is a P-matrix, then C can written as $C = A \oplus_k B$ so that A and B are both P-matrices.

Proof. The proof will use induction on k: the size of C_{22}. This result holds when $k = 1$, by Lemma 4.3. So assume the result is true for all $q \leqslant k - 1$. Let $q = k$, and suppose C is as indicated above. Let C_{11} be $m \times m$ and C_{33} be $r \times r$. Consider $C(m + k)$. $C(m + k)$ is a P-matrix in the same form as C, but with $q = k - 1$. Hence, by the induction hypothesis, $C(m + k)$ can be written as $C(m + k) = A' \oplus_{k-1} B'$, in which A' and B' are both P-matrices. Similarly, by induction, $C(\{m + 1, m + 2, \ldots, m + k - 1\})$ can be written as $C(\{m + 1, m + 2, \ldots, m + k - 1\}) = A'' \oplus_1 B''$, in which A'' and B'' are both P-matrices. Thus far we have decomposed C_{22} as follows: $C_{22}(k) = A'_{22} + B'_{22}$, in which A'_{22} and B'_{22} are both P-matrices, and $C_{22}[k] = \alpha'' + \beta''$, where α'' and β'' are both positive real numbers. Then

$$A' = \begin{bmatrix} C_{11} & C_{12}(\phi|\{k\}) \\ C_{21}(\{k\}|\phi) & A'_{22} \end{bmatrix} \quad \text{and} \quad B' = \begin{bmatrix} B'_{22} & C_{23}(\{k\}|\phi) \\ C_{32}(\phi|\{k\}) & C_{33} \end{bmatrix}$$

$$A'' = \begin{bmatrix} C_{11} & C_{12}[M|\{k\}] \\ C_{21}[\{k\}|M] & \alpha'' \end{bmatrix} \quad \text{and} \quad B'' = \begin{bmatrix} \beta'' & C_{23}[\{k\}|R] \\ C_{32}[R|\{k\}] & C_{33} \end{bmatrix},$$

where $M = \{1, 2, \ldots, m\}$ and $R = \{1, 2, \ldots, r\}$. Distribute the remaining entries of C_{22}, namely $C_{22}[\{k\}|\{1, 2, \ldots, k-1\}]$ and $C_{22}[\{1, 2, \ldots, k-1\}|\{k\}]$ as follows, so that $C = \tilde{A} \oplus_k \tilde{B}$, for

$$\tilde{A} = \begin{bmatrix} C_{11} & C_{12}(\phi|\{k\}) & C_{12}[M|\{k\}] \\ C_{21}(\{k\}|\phi) & A'_{22} & C_{22}[\{1, 2, \ldots, k-1\}|\{k\}] \\ C_{21}[\{k\}|M] & 0 & \alpha'' \end{bmatrix}$$

and

$$\tilde{B} = \begin{bmatrix} B'_{22} & 0 & C_{23}(\{k\}|\phi) \\ C_{22}[\{k\}|\{1, 2, \ldots, k-1\}] & \beta'' & C_{23}[\{k\}|R] \\ C_{32}(\phi|\{k\}) & C_{32}[R|\{k\}] & C_{33} \end{bmatrix}.$$

Notice that by construction both $\tilde{A}$ and $\tilde{B}$ satisfy the hypotheses of Lemma 2.6. Thus for a fixed nonzero $k \times k$ skew-symmetric matrix S with all off-diagonal entries nonzero in the last row and column and all other off-diagonal entries zero, there exists $T_{\tilde{A}} > 0$ so that, for all $t \geqslant T_{\tilde{A}}$,

$$\tilde{A} + \begin{bmatrix} 0 & 0 \\ 0 & tS \end{bmatrix},$$

is a P-matrix. Similarly, there exists a $T_{\tilde{B}} > 0$ so that, for all $t \geqslant T_{\tilde{B}}$,

$$\tilde{B} + \begin{bmatrix} t(-S) & 0 \\ 0 & 0 \end{bmatrix},$$

is a P-matrix. Let $T = \max\{T_{\tilde{A}}, T_{\tilde{B}}\}$. Then, for all $t \geqslant T$,

$$A = \tilde{A} + \begin{bmatrix} 0 & 0 \\ 0 & tS \end{bmatrix} \quad \text{and} \quad B = \tilde{B} + \begin{bmatrix} t(-S) & 0 \\ 0 & 0 \end{bmatrix},$$

are both P-matrices. Furthermore $C = A \oplus_k B$. This completes the proof. □

Finally, we consider the case $k = 1$, for the class P_0.

Theorem 4.6. *Let*

$$C = \begin{bmatrix} C_{11} & c_{12} & 0 \\ c_{21} & c_{22} & c_{23} \\ 0 & c_{32} & C_{33} \end{bmatrix} \in M_n(\mathbb{R})$$

be an $n \times n$ *matrix with* C_{11} *and* C_{33} *square, and with* c_{22} 1×1. *Then* C *is* P_0 *if and only if* C *may be written as* $C = A \oplus_1 B$, *in which both* A *and* B *are* P_0.

Proof. The sufficiency follows easily from Eq. (1) and the inheritance property for principal submatrices of P_0-matrices (as in the case for P-matrices). For the converse, we consider two cases: $c_{22} = 0$; or $c_{22} > 0$. If $c_{22} = 0$, then

$$C = \begin{bmatrix} C_{11} & c_{12} & 0 \\ c_{21} & 0 & c_{23} \\ 0 & c_{32} & C_{33} \end{bmatrix} = \begin{bmatrix} C_{11} & c_{12} \\ c_{21} & 0 \end{bmatrix} \oplus_1 \begin{bmatrix} 0 & c_{23} \\ c_{32} & C_{33} \end{bmatrix}.$$

Each of the two matrices on the right is easily seen to be P_0. Hence, suppose $c_{22} > 0$. Let $a_{22} \in [0, c_{22}]$, and define $b_{22} = c_{22} - a_{22}$. Now let

$$\hat{A} = \begin{bmatrix} C_{11} & c_{12} \\ c_{21} & a_{22} \end{bmatrix} \quad \text{and} \quad \hat{B} = \begin{bmatrix} b_{22} & c_{23} \\ c_{32} & C_{33} \end{bmatrix}.$$

Then $\hat{A} \oplus_1 \hat{B} = C$. We need to show there exists a choice for a_{22} so that $\hat{A}$ and $\hat{B}$ are P_0. Suppose, for simplicity of notation, that $\hat{A} \in M_m(\mathbb{R})$ and $\hat{B} \in M_r(\mathbb{R})$. Further, let

$$X = \{\alpha \subseteq \{1, 2, \ldots, m-1\} : \hat{A}[\alpha] \text{ is nonsingular}\},$$

and let

$$Y = \{\beta \subseteq \{2, 3 \ldots, r\} : \hat{B}[\beta] \text{ is nonsingular}\}.$$

Following similar analysis as in the proof of Lemma 4.3, it follows that a_{22} can be chosen so that

$$\max\{0, \max_{\alpha \in X}\{c_{21}[\alpha]\hat{A}[\alpha]^{-1}c_{12}[\alpha]\}\} \leqslant a_{22} \leqslant \min\{c_{22}, \min_{\beta \in Y}\{c_{22} - c_{23}[\beta]\hat{B}[\beta]^{-1}c_{32}[\beta]\}\}. \tag{3}$$

Futhermore, if a_{22} is chosen so that Eq. (3) holds, then $\hat{A}$ and $\hat{B}$ will be P_0. We will show that $\hat{A}$ is P_0; similar analysis shows $\hat{B}$ is P_0. Note that in order to show $\hat{A}$ is P_0 it is enough to show det $(\hat{A}[\alpha \cup \{m\}]) \geqslant 0$, for every $\alpha \subseteq \{1, 2, \ldots, m-1\}$. The remaining principal minors are nonnegative by assumption.

There are two cases to consider; $\hat{A}[\alpha]$ is singular or $\hat{A}[\alpha]$ is nonsingular. Firstly, suppose $\hat{A}[\alpha]$ is singular. Then

$$\det(\hat{A}[\alpha \cup \{m\}]) = \begin{bmatrix} \hat{A}[\alpha] & c_{12}[\alpha] \\ c_{21}[\alpha] & a_{22} \end{bmatrix} = \begin{bmatrix} \hat{A}[\alpha] & c_{12}[\alpha] \\ c_{21}[\alpha] & c_{22} \end{bmatrix} \geqslant 0,$$

since C is P_0. On the other hand, if $\hat{A}[\alpha]$ is nonsingular, then

$$\det(\hat{A}[\alpha \cup \{m\}]) = \det \hat{A}[\alpha](a_{22} - c_{21}[\alpha]\hat{A}[\alpha]^{-1}c_{12}[\alpha]) \geqslant 0,$$

by Eq. (3). Therefore $\hat{A}$ is P_0, similarly it follows that $\hat{B}$ is P_0 and by construction $\hat{A} \oplus_1 \hat{B} = C$. This completes the proof. □

When $k > 1$, it follows from Example 4.2, that the answer to question (III) is, in general, negative for the class P_0. We have not resolved question (IV) for the class P_0. However, in the special case in which

$$C = \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{21} & C_{22} & C_{23} \\ 0 & C_{32} & C_{33} \end{bmatrix} \in M_n(\mathbb{R}),$$

is P_0, and both C_{11} and C_{33} are P-matrices, then the same argument used to prove Theorem 4.5 may be applied to yield an affirmative answer to question (IV).

5. Totally nonnegative matrices

As a final topic we consider the class TN. We begin with the case $k = 1$. The following simple, but useful observation will be used throughout the proof of the next theorem. Suppose A is an $n \times n$ totally nonnegative matrix, and let E_{ij} denote the $n \times n$ standard basis matrix in which the (i,j)th entry is a 1 and all remaining entries are zero. Then $A + tE_{11}$ and $A + tE_{nn}$ are totally nonnegative matrices for all $t \geqslant 0$. The proof of this observation is routine and is omitted. It is noteworthy to recall the statement of *Fischer's inequality* (see [7, p. 117]). An $n \times n$ P_0-matrix A satisfies Fischer's inequality, if

$$\det A[\alpha \cup \beta] \leqslant \det A[\alpha] \det A[\beta],$$

for all $\alpha, \beta \subseteq N$ with $\alpha \cap \beta = \emptyset$. It is well known (see [7, p. 133]) that any totally nonnegative matrix satisfies Fischer's inequality. The next result answers questions (I) and (II) for the class TN.

Theorem 5.1. *Let*

$$C = \begin{bmatrix} C_{11} & c_{12} & 0 \\ c_{21} & c_{22} & c_{23} \\ 0 & c_{32} & C_{33} \end{bmatrix} \in M_n(\mathbb{R})$$

be an $n \times n$ matrix with C_{11} and C_{33} square, and with c_{22} 1×1. Then C is TN if and only if C may be written as $C = A \oplus_1 B$, in which both A and B are TN.

Proof. First, suppose both A and B are TN, and let

$$A = \begin{bmatrix} A_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in M_m(\mathbb{R}) \text{ and } B = \begin{bmatrix} b_{22} & b_{23} \\ b_{32} & B_{33} \end{bmatrix} \in M_r(\mathbb{R}),$$

where a_{22} and b_{22} are both 1×1. It has already been shown, in the case of P_0-matrices, that all the principal minors of $A \oplus_1 B$ are nonnegative. Let $C = A \oplus_1 B$, and let $C[\alpha|\beta]$ be any square submatrix of C. Let $\alpha_1 = \alpha \cap \{1, 2, \ldots, m-1\}$, $\alpha_2 = \alpha \cap \{m+1, m+2, \ldots, n\}$, $\beta_1 = \beta \cap \{1, 2, \ldots, m-1\}$ and $\beta_2 = \beta \cap \{m+1, m+2, \ldots, n\}$. Further, we can assume $\alpha_1, \alpha_2, \beta_1$ and β_2 are all nonempty, otherwise it is straightforward to verify that $\det C[\alpha|\beta] \geqslant 0$. Suppose $m \notin \alpha \cap \beta$. Then either $m \notin \alpha$ and $m \notin \beta$ or, without loss of generality, $m \in \alpha$ and $m \notin \beta$. First assume $m \notin \alpha$ and $m \notin \beta$. Then

$$C[\alpha|\beta] = \begin{bmatrix} A_{11}[\alpha_1|\beta_1] & 0 \\ 0 & B_{33}[\alpha_2|\beta_2] \end{bmatrix}.$$

If $|\alpha_1| = |\beta_1|$, then $C[\alpha|\beta]$ is a direct sum of TN matrices, and hence is TN. Otherwise $|\alpha_1| \neq |\beta_1|$, and without loss of generality, assume $|\alpha_1| > |\beta_1|$ (the case $|\alpha_1| < |\beta_1|$ follows by symmetry). Therefore the size of the larger zero block is $|\alpha_1| \times |\beta_2|$. Futhermore,

$$|\alpha_1| + |\beta_2| \geqslant |\beta_1| + |\beta_2| + 1 = |\beta| + 1 = |\alpha| + 1,$$

hence $\det C[\alpha|\beta] = 0$ (see Ref. [9], p. 55). Now assume $m \in \alpha$ and $m \notin \beta$. Then

$$C[\alpha|\beta] = \begin{bmatrix} A_{11}[\alpha_1|\beta_1] & 0 \\ a_{21}[\beta_1] & b_{23}[\beta_2] \\ 0 & B_{33}[\alpha_2|\beta_2] \end{bmatrix}.$$

If $|\alpha_1| = |\beta_1|$, then $|\alpha_2| + 1 = |\beta_2|$. Hence $C[\alpha|\beta]$ is block triangular and it follows that $\det C[\alpha|\beta] \geqslant 0$. Otherwise, $|\alpha_1| \neq |\beta_1|$. If $|\alpha_1| > |\beta_1|$, then $\det C[\alpha|\beta] = 0$ (see Ref. [9], p. 55]), and if $|\alpha_1| < |\beta_1|$, then either $C[\alpha|\beta]$ is block triangular or $C[\alpha|\beta]$ is singular. Thus suppose $m \in \alpha \cap \beta$. Again there are two cases to consider. Suppose $|\alpha_1| = |\beta_1|$. Then $|\alpha_2| = |\beta_2|$, and $\det C[\alpha|\beta] \geqslant 0$ follows from Eq. (1). Otherwise suppose $|\alpha_1| \neq |\beta_1|$, and without loss of generality, assume $|\alpha_1| > |\beta_1|$ (the case $|\alpha_1| < |\beta_1|$ follows by symmetry). The order of the larger zero block is $|\alpha_1| \times |\beta_2|$, and

$$|\alpha_1| + |\beta_2| \geqslant |\beta_1| + |\beta_2| + 1 = |\beta| = |\alpha|.$$

If $|\alpha_1| + |\beta_2| > |\beta|$, then, as before, det $C[\alpha|\beta] = 0$, see [9, p. 55]. So assume $|\alpha_1| + |\beta_2| = |\beta|$, in which case, it follows that $C[\alpha|\beta]$ is a block triangular matrix with the diagonal blocks being TN, hence det $C[\alpha|\beta] \geqslant 0$.

Conversely, suppose C is TN. Since

$$\begin{bmatrix} C_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \text{ and } \begin{bmatrix} c_{22} & c_{23} \\ c_{32} & C_{33} \end{bmatrix},$$

are TN, choose $\hat{a}_{22} \geqslant 0$ and $\hat{b}_{22} \geqslant 0$ as small as possible so that

$$\hat{A} = \begin{bmatrix} C_{11} & c_{12} \\ c_{21} & \hat{a}_{22} \end{bmatrix} \quad \text{and} \quad \hat{B} = \begin{bmatrix} \hat{b}_{22} & c_{23} \\ c_{32} & C_{33} \end{bmatrix},$$

are both TN. Assume $\hat{A} \in M_m(\mathbb{R})$ and $\hat{B} \in M_r(\mathbb{R})$. In what follows we are only concerned with square submatrices of $\hat{A}$ and $\hat{B}$. Let

$$\Gamma = \{\hat{A}[\alpha|\beta] : m \in \alpha, \beta \subseteq \{1, 2, \ldots, m\} \text{ and } \hat{A}[\alpha|\beta] \text{ is singular}\},$$

and let

$$\Lambda = \{\hat{B}[\gamma|\delta] \colon 1 \in \gamma, \delta \subseteq \{1, 2 \ldots, r\} \text{ and } \hat{B}[\gamma|\delta] \text{ is singular}\}.$$

Observe that Γ (similarly Λ) is nonempty. Since if Γ was empty, then for every $m \in \alpha, \beta \subseteq \{1, 2, \ldots, m\}$, det $\hat{A}[\alpha|\beta] > 0$. Then by continuity we may decrease $\hat{a}_{22}$, while $\hat{A}$ remains TN, which contradicts the minimality of $\hat{a}_{22}$. Therefore Γ is nonempty and similarly so is Λ.

Suppose for the moment that $\hat{a}_{22} + \hat{b}_{22} \leqslant c_{22}$. Then increase $\hat{a}_{22}$ to a_{22} and increase $\hat{b}_{22}$ to b_{22} so that $a_{22} + b_{22} = c_{22}$, and let

$$A = \begin{bmatrix} C_{11} & c_{12} \\ c_{21} & a_{22} \end{bmatrix} \text{ and } B = \begin{bmatrix} b_{22} & c_{23} \\ c_{32} & C_{33} \end{bmatrix}.$$

By the observation preceeding Theorem 5.1, A and B are TN, and $C = A \oplus_1 B$. Thus, if we can show $\hat{a}_{22} + \hat{b}_{22} \leqslant c_{22}$, then the proof is complete. So suppose $\hat{a}_{22} + \hat{b}_{22} > c_{22}$. Then one of two possibilities can occur. Either $\hat{A}[\alpha - \{m\}|\beta - \{m\}]$ is singular, for every $\hat{A}[\alpha|\beta] \in \Gamma$, or there exists $\hat{A}[\alpha_0|\beta_0] \in \Gamma$ such that $\hat{A}[\alpha_0 - \{m\}|\beta_0 - \{m\}]$ is nonsingular. Suppose that $\hat{A}[\alpha - \{m\}|\beta - \{m\}]$ is singular, for every $\hat{A}[\alpha|\beta] \in \Gamma$. In this case, each such det $\hat{A}[\alpha|\beta]$ does not depend on $\hat{a}_{22}$, and hence $\hat{a}_{22}$ may be decreased without affecting det $\hat{A}[\alpha|\beta]$. Also, as previously noted, if $m \in \alpha', \beta' \subseteq \{1, 2, \ldots, m\}$ and det $\hat{A}[\alpha'|\beta'] > 0$, then $\hat{a}_{22}$ may be decreased in this case. However, this contradicts the minimality of $\hat{a}_{22}$. Thus there exists $\hat{A}[\alpha_0|\beta_0] \in \Gamma$ such that $\hat{A}[\alpha_0 - \{m\}|\beta_0 - \{m\}]$ is nonsingular. Similar arguments also show that there exists $\hat{B}[\gamma_0|\delta_0] \in \Lambda$ such that $\hat{B}[\gamma_0 - \{1\}|\delta_0 - \{1\}]$ is nonsingular. Furthermore, if $\hat{a}_{22}$ is decreased, then det $\hat{A}[\alpha_0|\beta_0] < 0$. Since $\hat{a}_{22} + \hat{b}_{22} > c_{22}$, decrease $\hat{a}_{22}$ to a'_{22} such that $a'_{22} + \hat{b}_{22} = c_{22}$. Then

$$C' = \begin{bmatrix} \hat{A}[\alpha_0 - \{m\}|\beta_0 - \{m\}] & c_{12}[\alpha_0 - \{m\}] \\ c_{21}[\beta_0 - \{m\}] & a'_{22} \end{bmatrix} \oplus_1 \hat{B}[\gamma_0|\delta_0],$$

is a submatrix of C. However, $0 \leqslant \det C' < 0$, by Eq. (1). This is a contradiction, hence $\hat{a}_{22} + \hat{b}_{22} \leqslant c_{22}$, which completes the proof. □

It is easy to verify that the class TN is not closed under addition, hence it follows that the answer to question (III) is negative. However, as we noted in other cases, even more can go wrong.

Example 5.2.

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then A and B are both totally nonnegative matrices. However,

$$A \oplus_2 B = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

is not a totally nonnegative matrix since $\det(A \oplus_2 B) = -2$. Note that the sum in the overlapping positions is a totally nonnegative matrix.

To address question (IV), we will make strong use of the fact [10] that TN matrices are LU factorable into TN matrices L and U, but we need rather more than what is known, see also [11]. When the TN matix C has our form

$$C = \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{21} & C_{22} & C_{23} \\ 0 & C_{32} & C_{33} \end{bmatrix},$$

we need that, in addition to being lower and upper triangular, respectively, the (correspondingly partitioned) L has its 3,1 block and U its 1,3 block 0. In case C_{11} is invertible, a simple partitioned calculation reveals that this must occur in any LU factorization of C (since L_{11} and U_{11} will be invertible). However, when C_{11} is singular (which implies by Fischer's inequality that C is singular), there can occur TN matrices C of our form that have TN LU factorizations with positive entries in the L_{31} or U_{13} blocks. Fortunately, though, LU factorizations will be highly nonunique in this case, and there will always exist ones of the desired form. Thus, an auxiliary assumption that C_{11} is invertible would avoid the need for Lemma 5.3 in the proof of Theorem 5.4, but this somewhat specialized lemma (perhaps of independent interest, though it requires an elaborate proof) shows that such an auxiliary assumption is unnecessary.

For simplicity of notation (see also [10]), we let

$$G_2 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix},$$

and let $F_r(c)$, $c \geqslant 0$, denote the following $r \times r$ matrix,

$$F_r(c) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0_{r-2} & 0 \\ c & 0 & 1 \end{bmatrix}.$$

Observe, that both of the above matrices are TN. We also note that the property TN is not in general preserved under row and column operations. However, it follows (see, for example, [12]), that if $A = [a_{ij}]$ is TN and say $a_{1j}, a_{1,j+1} > 0$, then the matrix A' obtained from A by subtracting $a_{1,j+1}/a_{1j}$ times column j from column $j+1$ is TN. That is $A = A'U$, where U is the upper triangular TN matrix given by $U = I_{j-1} \oplus F_2^{\mathrm{T}}(a_{1,j+1}/a_{1j}) \oplus I_{n-j-1}$. This notion was generalized in Ref. [10]. Suppose $A = [a_{ij}]$ is TN, $a_{1j} > 0$, and for some $t > j$, $a_{1t} > 0$, but $a_{1k} = 0$ for all $j < k < t$. Then, since A is TN it follows that $a_{ik} = 0$ for $1 \leqslant i \leqslant n$ and $j < k < t$. Let A' be obtained from A by subtracting a_{1t}/a_{1j} times column j from column t. Thus $A = A'U$, where U is the upper triangular TN matrix given by $U = I_{j-1} \oplus F^{\mathrm{T}}_{t-j+1}(a_{1t}/a_{1j}) \oplus I_{n-t}$. It is shown in Ref. [10] that the matrix A' is TN.

Lemma 5.3. *Let*

$$C = \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{21} & C_{22} & C_{23} \\ 0 & C_{32} & C_{33} \end{bmatrix},$$

in which C_{22} and C_{33} are square and C_{11} is $m \times m$ ($m \geqslant 1$). Suppose C is TN. Then

$$C = LU = \begin{bmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ 0 & L_{32} & L_{33} \end{bmatrix} \cdot \begin{bmatrix} U_{11} & U_{12} & 0 \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{bmatrix},$$

in which L and U (partitioned conformally with C) are both TN.

Proof. First, we note that the algorithm given in Ref. [10] applied to C, always produces a U, with $U_{13} = 0$. The proof of the lemma will use induction on m, the size of C_{11}. Suppose $m = 1$, and consider two cases; $C_{11} > 0$ (C_{11} is 1×1), or $C_{11} = 0$. If $C_{11} > 0$, then C_{11} is nonsingular and we are done by previous observations. Thus suppose $C_{11} = 0$. Since C is TN, either row or column 1 must be zero. Assume, without loss of generality, that $C_{12} = 0$; otherwise consider C^{T}. If $C_{21} = 0$ as well, then by the comments on the previous page C has a desired LU-factorization. Therefore assume $C_{21} \neq 0$. By the algorithm given in Ref. [10] C may be written as,

$$C = \begin{bmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{bmatrix} \cdot \begin{bmatrix} U_{11} & U_{12} & 0 \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{bmatrix}.$$

Since $C_{21} \neq 0$, it follows that $U_{11} \neq 0$. However, $L_{31}U_{11} = 0$, hence $L_{31} = 0$. Now assume the result is true for all such C, with C_{11} $l \times l$, $1 \leqslant l \leqslant m-1$. Suppose $C = [c_{ij}]$ is as above, with C_{11} $m \times m$. Again we consider two cases; $c_{11} = 0$ and $c_{11} > 0$. Suppose $c_{11} = 0$, and assume, without loss of generality, column 1 of C is zero. Let j be the smallest integer (larger than 1) for which $c_{1j} > 0$. If no such j exists, then the result follows easily by induction. Since C is TN it follows that $c_{it} = 0$, for all i, t such that $t < j$. Applying column operations like those discussed before the lemma, C may be written as

$$C = L'U' = \left[\begin{array}{ccc|c} 0 & 0 & c_{1j} & 0 \\ \hline 0 & 0 & c_{2j} & \\ \vdots & \vdots & \vdots & C' \\ 0 & 0 & c_{nj} & \end{array}\right] \cdot U',$$

(C' is $(n-1) \times (n-j)$) where L' and U' are TN. Since the first $j-1$ columns of C are all 0, observe that L' may be written as

$$L' = \left[\begin{array}{ccc|c} c_{1j} & 0 & 0 & 0 \\ \hline 0 & 0 & c_{2j} & \\ \vdots & \vdots & \vdots & C' \\ 0 & 0 & c_{nj} & \end{array}\right] \cdot U^{(1,2)}U^{(2,3)} \cdots U^{(j-1,j)},$$

in which $U^{(r,r+1)} = I_{r-1} \oplus G_2 \oplus I_{n-r-1}$. Note, when $r = 1$ we ignore the first summand. It is easy to verify that $U^{(r,r+1)}$ is TN. Let

$$L'' = \left[\begin{array}{ccc|c} c_{1j} & 0 & 0 & 0 \\ \hline 0 & 0 & c_{2j} & \\ \vdots & \vdots & \vdots & C' \\ 0 & 0 & c_{nj} & \end{array}\right]$$

and let $U'' = U^{(1,2)}U^{(2,3)} \cdots U^{(j-1,j)}U'$. Then U'' is TN since U'' is a product of TN matrices and L'' is TN since L' is TN, and $C = L''U''$. Now let

$$C'' = \left[\begin{array}{cc|c|} 0 & c_{2j} & \\ \vdots & \vdots & C' \\ 0 & c_{nj} & \end{array}\right]$$

(C'' is $(n-1) \times (n-1)$). By induction, $C'' = \bar{L}\bar{U}$, in which $\bar{L}$ and $\bar{U}$ are TN, and with $\bar{L}_{31} = 0$ and $\bar{U}_{13} = 0$. Then

$$C = \begin{bmatrix} c_{1j} & 0 \\ 0 & \bar{L} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & \bar{U} \end{bmatrix} \cdot U'',$$

is a desired LU factorization. Finally, suppose $c_{11} > 0$, and assume column 1 of C is equal to

$$[c_{11}, c_{21}, \ldots, c_{l1}, 0, \ldots, 0, c_{j1}, 0, \ldots, 0]^{\mathrm{T}},$$

where $c_{l1}, c_{j1} > 0$ and $1 \leqslant l \leqslant j-1$. Since C is TN and $c_{i1} = 0$, for $l < i < j$, $c_{it} = 0$, for $l < i < j$ and $1 \leqslant t \leqslant n$. Observe that if $L^{(l,j)} = I_{l-1} \oplus F_{j-l+1}(c_{j1}/c_{l1}) \oplus I_{n-j}$, then $L^{(l,j)}$ is TN. Furthermore, C may be written as $C = L^{(l,j)}C'$, in which the first column of C' is equal to

$$[c_{11}, c_{21}, \ldots, c_{l1}, 0, \ldots, 0]^{\mathrm{T}}.$$

Moreover, it follows from a result in Ref. [10], that C' is TN. Applying similar row operations we obtain $C = L'\tilde{C}$, in which L' and $\tilde{C}$ are both TN and with the first column of $\tilde{C}$ given by $[c_{11}, 0, \ldots, 0]^{\mathrm{T}}$. Note that L' is equal to a product of lower triangular TN matrices each of the form $L^{(s,t)}$. Applying the first step of the algorithm given in Ref. [10] we may write $\tilde{C}$ as

$$\tilde{C} = \tilde{L}U' = \begin{bmatrix} c_{11} & 0 \\ 0 & C'' \end{bmatrix} \cdot U',$$

in which $\tilde{L}$ and U' are both TN. By induction we may write $C'' = L''U''$, in which L'' and U'' are TN, and with $L''_{31} = 0$ and $U''_{13} = 0$. Thus we may write C as

$$C = L' \begin{bmatrix} c_{11} & 0 \\ 0 & L'' \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & U'' \end{bmatrix} U'.$$

which is a desired decomposition. This completes the proof. □

We are now in a position to prove an affirmative answer to question (IV) for the class TN.

Theorem 5.4. *Let*

$$C = \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{21} & C_{22} & C_{23} \\ 0 & C_{32} & C_{33} \end{bmatrix},$$

in which C_{11} and C_{33} are square, and with C_{22} $k \times k$. If C is TN, then C can be written as $C = A \oplus_k B$ so that A and B are TN.

Proof. By Lemma 5.3, $C = LU$, in which both L and U are TN, and $L_{31} = 0$ and $U_{13} = 0$. Then it is easy to check that

$$C = LU = \begin{bmatrix} L_{11}U_{11} & L_{11}U_{12} & 0 \\ L_{21}U_{11} & L_{22}U_{22} + L_{21}U_{12} & L_{22}U_{23} \\ 0 & L_{32}U_{22} & L_{33}U_{33} + L_{32}U_{23} \end{bmatrix}.$$

Hence C can be written as

$$C = \begin{bmatrix} L_{11}U_{11} & L_{11}U_{12} & 0 \\ L_{21}U_{11} & L_{21}U_{12} & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & L_{22}U_{22} & L_{22}U_{23} \\ 0 & L_{32}U_{22} & L_{33}U_{33} + L_{32}U_{23} \end{bmatrix}.$$

Notice that if

$$A = \begin{bmatrix} L_{11}U_{11} & L_{11}U_{12} \\ L_{21}U_{11} & L_{21}U_{12} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \cdot \begin{bmatrix} U_{11} & U_{12} \\ 0 & 0 \end{bmatrix},$$

and

$$B = \begin{bmatrix} L_{22}U_{22} & L_{22}U_{23} \\ L_{32}U_{22} & L_{33}U_{33} + L_{32}U_{23} \end{bmatrix} = \begin{bmatrix} L_{22} & 0 \\ L_{32} & L_{33} \end{bmatrix} \cdot \begin{bmatrix} U_{22} & U_{23} \\ 0 & U_{33} \end{bmatrix},$$

then $C = A \oplus_k B$. Each of the four matrices on the right is easily seen to be TN because L and U are, and it follows from the multiplicative closure of the class TN that both A and B are TN, which completes the proof. □

For each class and for each question ((I)–(IV)) we indicate when the answer to the question is true (T), or false (F). The asterisk indicates a partial affirmative answer for the class P_0 (see discussion after Theorem 4.6).

Acknowledgements

The authors are grateful to Professor Jane Day for very helpful discussions which led to Lemma 5.3.

Table 1
Summary of the answers to questions (I)–(IV) for each of the classes considered

Class	Question (I)	(II)	(III)	(IV)
PSD	T	T	T	T
PD	T	T	T	T
M	T	T	F	T
TN	T	T	F	T
P	T	T	F	T
CP	T	T	T	T
DN	T	T	T	F
SM	T	T	T	T
P_0	T	T	F	T

References

[1] J.H. Drew, C.R. Johnson, The completely positive and doubly nonnegative completion problems, Linear and Multilinear Alg. 44 (1998) 85–92.
[2] R. Grone, C.R. Johnson, E. Sa, H. Wolkowicz, Positive definite completions of partial hermitian matrices, Linear Algebra Appl. 58 (1984) 109–124.
[3] C.R. Johnson, Matrix completion problems: A Survey, Proceedings of Symposia in Applied Mathematics 40 (1990) 171–198.
[4] C.R. Johnson, R.L. Smith, The completion problem for M-matrices and inverse M-matrices, Linear algebra Appl. 241–243 (1996) 655–667.
[5] J.H. Drew, C.R. Johnson, The no long odd cycle theorem for completely positive matrices, The IMA Volumes in Mathematics and its Applications, Springer, New York, vol. 76, 1996, pp. 103–115.
[6] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, New York, 1985.
[7] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, New York, 1991.
[8] C.R. Johnson, B.K. Kroschel, The combinatorially symmetric P-matrix completion problem, Electronic Journal of Linear Algebra 1 (1996) 59–63.
[9] H.J. Ryser, Combinatorial Mathematics, The Mathematical Association of America, Rahway, 1963.
[10] C.W. Cryer, Some properties of totally positive matrices, Linear Algebra Appl. 15 (1976) 1–25.
[11] T. Ando, Totally positive matrices, Linear Algebra Appl. 90 (1987) 165–219.
[12] A. Whitney, A reduction theorem for totally positive matrices, J. Analyse Math. 2 (1952) 88–92.